Morbidity severity classifying routine consultations from English and Dutch general practice indicated physical health status

Kadam, U.T.; Schellevis, F.G.; van der Windt, D.A.W.M.; de Vet, H.C.W.; Bouter, L.M.; Croft, P.R.

published in
Journal of Clinical Epidemiology
2008

DOI (link to publisher)
10.1016/j.jclinepi.2007.05.014

document version
Publisher's PDF, also known as Version of record

ELSEVIER

Journal of Clinical Epidemiology 61 (2008) 386–393

Journal of Clinical Epidemiology

Morbidity severity classifying routine consultations from English and Dutch general practice indicated physical health status

U.T. Kadam[a,*], F.G. Schellevis[b,d], D.A.W.M. van der Windt[a,c], H.C.W. de Vet[c], L.M. Bouter[c], P.R. Croft[a]

[a]*Primary Care Sciences Research Centre, Keele University, Staffordshire, England ST5 5BG, United Kingdom*
[b]*NIVEL Netherlands Institute for Health Services Research, Postbus 1568, 3500 BN Utrecht, The Netherlands*
[c]*EMGO Institute, VU University Medical Center, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands*
[d]*Department of General Practice, VU University Medical Center, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands*

Accepted 10 May 2007

Abstract

Objective: To investigate the construct validity of morbidity severity scales based on routine consultation data by studying their associations with sociodemographic factors and physical health.

Study Design and Setting: Study participants were 11,232 English adults aged 50 years and over and 9,664 Dutch adults aged 18 years and over, and their consulting morbidity data in a 12-month period were linked to their physical health data. Consulters with any of 115 morbidities classified on four ordinal scales of severity (''chronicity,'' ''time course,'' ''health care use,'' and ''patient impact'') were compared to all other consulters.

Results: As hypothesized, in both countries, morbidity severity was associated with older age, female gender, more deprivation (all comparisons $P \leq 0.05$), and poor physical health (all trends $P < 0.001$). The estimated strengths of association of poor physical health with the highest severity category expressed as odds ratios, for each of the four scales, were 5.4 for life-threatening on the ''chronicity'' scale, 1.8 for time course, 2.8 for high health care use, and 3.7 for high patient impact.

Conclusions: Four scales of morbidity severity have been validated in English and Dutch settings, and they offer the potential to use simple routine consultation data as an indicator of physical health status in populations from general practice.

Keywords: Morbidity; General practice; Medical records; Health status indicators; Comorbidity; Classification

1. Introduction

In the United Kingdom [1] and the Netherlands, most of the population are registered with a general practitioner (GP), and in an average British practice there are an estimated 50,000 consultations per year [2]. Population-based consultation data provide an estimate of morbidity and have been used to study health needs relating to specific conditions [2,3], and to assess health care use as an outcome of primary care interventions [4,5]. The focus of attention in health care tends to be on arbitrarily defined chronic diseases, major life-threatening disorders such as cancer, or diseases that result in hospital admissions. Little attention has been paid to the much broader spectrum of morbidity which people present to primary care. The importance of this spectrum lies in the fact that it relates to the whole health experience of patients, which is inclusive of symptoms and chronic health states, and the need for the comprehensive health care which general practice attempts to address. In this spectrum, however, individual morbidities may vary in the extent to which they impair overall health, and it is possible that the cumulative effect of consultations for different morbidities over time may provide a more powerful and useful severity measure of individual health status than single and specific diagnostic labels, and yet could still be derived from routinely collected data. The practical usefulness of allocating morbidity to a ''severity'' measure from the GP viewpoint is that it might enable the grouping of consulters based on overall health need (of single and multiple morbidity consultations) and for planning health care delivery. An alternative approach is to use health-related quality-of-life questionnaires, but these can be cumbersome. Questionnaires provide estimates of health

* Corresponding author. Tel.: +44-1782-583-924; fax: +44-1782-583-911.

E-mail address: u.kadam@cphc.keele.ac.uk (U.T. Kadam).

0895-4356/08/$ – see front matter
doi: 10.1016/j.jclinepi.2007.05.014

status that are fixed in time unless repeated, but are also time and cost-intensive with concomitant ethical and patient recruitment issues.

The morbidity that patients' present is now routinely recorded by GPs in consultations, particularly in networks dedicated to collecting clinical data. Simple measures of contact frequency have already shown that patient's exhibit specific patterns of consultation [6]. Previous attempts to add a measure of morbidity severity to this basic picture have been based on additional assessment of severity in the consultation, focused on nonprimary care settings [7,8] or have used a restricted number of morbidities [9,10] that have not been validated fully. Such measurement of ''severity'' has taken two forms: (i) based on a priori classification for routinely collected morbidity data, that is, morbidity severity classified relative to one another [7,9,10] or (ii) based on classification of the severity of morbidity in each patient [8]. We have worked with a group of UK GPs and used their clinician constructs of morbidity severity to produce a classification of the first type. Because GPs see the fullest range of morbidities, they are perhaps best placed to provide clinical constructs that can order the individual morbidities according to likely health severity. So in the North Staffordshire (English) setting, four scales of morbidity severity based on routine consultation data were first developed with GPs. The objective of the work described in this paper was to validate these severity scales against two constructs: self-reported physical health status and variation with individuals' sociodemographic characteristics.

In the development of any epidemiological measure, it is important that face validation is combined with other tests of validity such as content, construct, and predictive validity. An additional important component is the external validity or generalizability of the tool in different populations' [11]. So in a two-step process, we first validated the scales in general practices from England and then repeated the same validation process in general practices from the Netherlands. Dutch general practice provides an appropriate external reference to English general practice, as they have similar health care systems with general practice as the main access point to health services [12]. The consulting populations selected for study in both countries had also participated in health surveys, thus enabling consultation morbidity data, health status, and sociodemographic data to be linked.

2. Methods

2.1. Design

In preliminary work in England, we first carried out focus group meetings to develop scales of morbidity severity that were then applied to a list of morbidities for classification by GPs using consensus methods. In the main study, we applied the severity classification to routine consultation data for a 12-month period from six English general practices that were linked to individual patient health survey data, and to consultation data for the same time period from general practices in the Netherlands that were also linked to individual patient data from the second Dutch National survey (DNS2).

2.2. Preliminary work: development of morbidity severity scales

In England, a focus group of six experienced GPs (each with an average of more than 15 years of experience in general practice) was recruited from the North Staffordshire General Practice Research Network (NSGPRN). The objective of the meetings that were carried out in four two-hour sessions was to develop ways of characterizing severity of morbidity seen and recorded in general practice based on their own clinical experience. They developed four separate scales of severity, representing dimensions and categories, which they regarded as important, to classify morbidities seen in general practice according to their likely impact on overall health. For each scale, the GP focus group created a priori ordinal ''severity'' categories as follows: (i) ''chronicity''—acute, acute-on-chronic, chronic, life-threatening, (ii) ''time course''—one-off, recurrent, progressive, and permanent, (iii) ''health care use''—low, medium, and high, (iv) ''patient impact'' on activities of daily living—low, medium, and high. So the a priori hypothesis of the GP focus group was that higher morbidity severity would indicate poorer overall health status. Although the viewpoint of the focus group was that the scales would overlap, nevertheless they would provide differing (multidimensional) clinical perspectives on measuring ''severity'' and address the spectrum of individual morbidities in general practice, some of which might be more classifiable on one scale than on another.

In a two-round consensus process, 44 GPs from NSGPRN classified a total of 188 Read-coded morbidities [13] according to these four scales, and a morbidity was classified to a severity category if there was agreement among 66% of the GPs or more. The 188 morbidities were selected from consulters aged 50 years and over in an anonymized English general practice database collected during 1 year [2]. The selection included 56 most prevalent, 114 randomly selected, and 18 selected on the basis of an earlier study of older patients. The 56 (that includes commonly screened conditions) with a prevalence of more than 0.34% will identify most of the cases providing the ''sensitivity'' in the measurement of morbidity severity. The remaining less prevalent morbidities provide a random allocation of patients to the severity groups and provide the ''specificity'' in the measurement. So from an epidemiological and a probability perspective, the case definition for severity will be driven by the commoner morbidities. This consensus project is described elsewhere [14].

2.3. Main study

For the validation study described here, the first objective was to test the construct of the four scales of severity using routine consultations from the English general practice population and then to test the same constructs and generalizability by application to another general practice consulters from another population—this time in the Netherlands. All phases in England had Research Ethics Committee approval, but this was not required for using the specific anonymized data from the DNS2 [15].

2.4. General practices

At the time of the study data collection (2001), the NSGPRN had 17 practices covering a wide range of socioeconomic groups, urban and rural populations, and over 70 GPs who had actively participated in studies of Read code use since 1996 [16]. The DNS2 included a 1-year survey of GP consultations in 104 general practices in the Netherlands (2001), comprising 195 GPs.

2.5. Consulting morbidity data

GPs in the Netherlands had used the ICPC-1 (International Classification for Primary Care) [17] to code consulting morbidities, whereas North Staffordshire GPs had used Read codes [13]. Of the original set of 188 morbidities, 115 classified morbidities were similar in the Read and ICPC-1 coding systems.

The classification of the 115 morbidities according to the severity scales was as follows: 85 (73%) by chronicity (46 acute, 11 acute-on-chronic, 21 chronic, and seven life-threatening); 87 (75%) by time course (37 one-off, 24 recurrent, 22 progressive, and four permanent); 76 (66%) by health care use (39 low, 29 medium, and eight high health care use); and 68 (59%) by patient impact (13 low, 34 medium, and 21 high impact). All 115 morbidities had a severity classification on at least one scale, but only 31 morbidities on all four scales, which means that each scale categorizes different subgroups of patients within the consulting population (see Supplementary Table on the journal's website at www.elsevier.com/locate/jclinepi). Of the excluded 73 Read codes that could not be mapped to ICPC-1, almost all were low prevalence morbidities based on English data. The classification for these were (i) 42 acute, one acute-on-chronic, 12 chronic, and six life-threatening; (ii) 37 one-off, nine recurrent, 10 progressive, two permanent; (iii) 29 low, nine medium, six high health care use; and (iv) seven low, 20 medium, and 11 high patient impact.

2.6. Population surveys

In two previous studies, all "older" adults aged 50 years and over registered with six North Staffordshire practices had been sent postal health surveys (total baseline sample of 20,133), and 14,670 (73%) patients had responded, of whom 11,232 consented to the review of their clinical records [18,19]. Information downloaded in clinical records included postcodes to determine deprivation status (Townsend score [20] based on 2001 national UK census data). The Townsend score uses data on housing quality, car ownership, and number of people in the household to produce a composite score of relative deprivation. In the postal surveys, the Short-Form Medical Outcomes Study questionnaires (SF-12 in three practices and SF-36 in three practices) had been used as a generic measure of health status [21,22]. The consultation data during a 12-month period were linked to the cross-sectional health surveys at the end point of that period.

In the Netherlands, an all-age nationally representative sample of 19,685 was randomly chosen for a detailed health interview in 2001, and 12,699 (65%) responded. Of this total, 9,664 were adults, with 5,173 "younger" patients (18–49 years) and 4,491 "older" patients (50 years and over). Instruments in the interview had included the SF-36, and individual monthly income data were used as measure of deprivation. Consultation and survey data collected concurrently over a 12-month period were linked.

The Short-Form (12 and 36) questionnaires are validated and widely used generic measures of health. The six questions (subscales: Physical functioning (two questions), Role-Physical (two), bodily pain (one), and general health (one)) common to both instruments were used to summarize the Physical Component Score (PCS) in both data sets, and using the SF Health Outcomes Scoring Software, any missing data were also imputed to obtain the most complete set of summary scores [23]. The overall PCS score was used as a global measure of the physical health status.

2.7. Group definitions

Cases were all patients who had consulted for at least one of the 115 morbidities during the 12-month period, and patients were classified on at least one of the four severity scales. The reference group was patients who had consulted for all other morbidity, that is, morbidity "undefined" by severity, and the same reference was used for all four scales.

The definitions apply to at least one consultation for a given morbidity in the study time period and do not include repeat consultations for the same morbidity. Classification by a morbidity severity category relates to at least one consultation in the study time period and does not include multiple morbidities classified by the same severity category.

2.8. Construct validation hypotheses

The validation process was carried out in two steps in existing data sets, as these offer the most independently collected data appropriate for examining associations with severity classification. First, internal validity [11] was tested in older English consulters, with the construct hypotheses that all four ordinal scales of morbidity severity would be

positively associated with older age, females, higher deprivation, and poorer physical health adjusting for the same sociodemographic factors. Current evidence indicates that adverse health states are more common in older age groups [24], females [25], and those from economically deprived groups [26]. Second, external validation was tested by the hypothesis that the strength of associations observed in older English consulters would be similar in older Dutch consulters (aged 50 years and over) and separately that the associations would show similar trends in younger Dutch consulters (18–49 years). These hypotheses tested each category within an ordinal scale and were not meant to provide insights into temporal relationships between morbidity severity and physical health.

3. Statistical analysis

For each of the four scales separately, patients were categorized by the most severe grading for which they had a morbidity consultation in the study year, and two analyses were performed.

First, associations were estimated, for each scale separately, between severity category and (i) age categorized in four 10-year bands and 80+ for older English and Dutch consulters, and for younger Dutch consulters aged 18–34 and 35–49 years, (ii) gender and (iii) deprivation status (Townsend data and Dutch income data were dichotomized into ''affluent'' and ''deprived'' scores). Chi-square tests were used to assess trends of association, for each scale separately, between morbidity severity and sociodemographic characteristics using a statistical significance level of 5%.

Second, physical health scores from the SF-12 questions were dichotomized into ''poor'' and ''good'' using the mean score for the English sample as a whole, and the mean scores for the older and younger Dutch samples, respectively. Associations between morbidity severity and physical health in each scale were estimated using unconditional logistic regression with 95% confidence intervals (CIs) to compare the odds of poor physical health for each severity category compared with the reference group. The trend in odds ratio (OR) within each scale was assessed on the basis of null hypothesis in linear trend of association. For construct validation, crude estimates are relevant. To account for differences in age, gender, and deprivation status between English and Dutch samples, we have also presented adjusted values. All analyses were performed using SPSS version 11.0 for Windows.

4. Results

In the 12-month study period, 9,003 (80.2%) of the 11,232 older English sample had consulted for at least one coded morbidity. In the Dutch sample of 9,664, 3,757 (83.7%) of the ''older'' patients and 3,996 (77.2%) of the ''younger'' patients had consulted.

4.1. Classifying consulters on the severity scales

Of English consulters, 1,428 (15.9%) were classified to the reference group (severity undefined), and 7,575 (84.1%) were classified on at least one of the four severity scales. For the 3,757 older Dutch consulters, the figures were 699 (18.6%) in the reference group and 3,058 (81.4%) categorized by at least one of the scales, and for the 3,996 young Dutch consulters the figures were 1,506 (37.7%) and 2,490 (62.3%), respectively. The older English and Dutch consulters classified by case and reference groups were similar for age and gender, but Dutch patients were relatively more deprived, had higher morbidity counts, and better physical health (Table 1).

4.2. Morbidity severity and sociodemographic factors

Within all four scales, morbidity severity showed significant positive trends with increasing age (all trend tests $P < 0.001$) and higher socioeconomic deprivation (all trend tests $P < 0.01$ (English) and $P < 0.001$ (Dutch)) (Table 2). Severity was higher in females on three scales in English consulters ($P \leqslant 0.05$), the exception being health care use, and on all scales in Dutch consulters ($P < 0.05$ for chronicity and $P < 0.001$ for the other three scales).

4.3. Morbidity severity and physical health: older English consulters

On three scales (chronicity, health care use, and patient impact), severity was associated (trend tests $P < 0.001$) with poor physical health compared to the reference group (Table 3). This was also true for time course except that the OR for ''progressive'' (4.5) was higher than that for ''permanent'' category (2.2) (on the ordinal scale, the latter category had been predefined as more severe than the former). The ORs were marginally reduced when adjusting for age, gender, and deprivation.

4.4. Morbidity severity and physical health: older and younger Dutch consulters

Overall, results were similar in the different study populations with few exceptions (Table 3). Estimates for older Dutch consulters were similar in strength of associations to older English consulters on all scales. Younger Dutch consulters generally showed similarly significant trends in associations (minimum $P \leqslant 0.05$), although the ORs were smaller when compared to the older consulters.

4.5. Nonresponse bias

In the English study, responders were more likely to be females ($P < 0.001$), older ($P < 0.001$), and affluent ($P < 0.001$) than nonresponders, whereas consenters to record review tended to be relatively younger ($P < 0.001$), male ($P < 0.001$), and affluent ($P < 0.001$) than nonconsenters. There was no statistical difference between record

Table 1
Characteristics of the study populations classified by four severity scales

Consulters	Classified by severity	Severity scales	Age (yr) Mean (SD)	Gender % M:F	Deprivation % affluent:deprived	Morbidity count Mean (SD)	Physical health Mean (SD)
English (50+ yr, $n = 9,003$)	No ($n = 1,428$)	Reference[a] ($n = 1,428$)	63.7 (9.8)	48:52	62:38	1.5 (0.90)	43.2 (11.8)
	Yes ($n = 7,575$)	Chronicity ($n = 6,408$)	66.8 (9.8)	43:57	54:46	3.6 (2.4)	38.2 (12.0)
		Time course ($n = 6,451$)	66.7 (10.1)	43:57	54:46	3.6 (2.4)	38.3 (12.0)
		Health care use ($n = 6,043$)	66.6 (10.1)	42:57	55:45	3.7 (2.4)	38.3 (12.0)
		Patient impact ($n = 6,007$)	66.5 (10.1)	44:56	55:45	3.7 (2.4)	37.7 (11.9)
Dutch (50+ yr, $n = 3,757$)	No ($n = 699$)	Reference[a] ($n = 699$)	62.0 (9.7)	46:54	43:57	1.9 (1.3)	48.8 (8.6)
	Yes ($n = 3,058$)	Chronicity ($n = 2,653$)	66.1 (10.6)	44:56	34:66	6.3 (3.7)	43.5 (10.8)
		Time course ($n = 2,671$)	66.1 (10.6)	43:57	34:66	6.3 (3.7)	43.5 (10.8)
		Health care use ($n = 2,624$)	66.1 (10.6)	42:58	35:65	6.3 (3.7)	43.5 (10.8)
		Patient impact ($n = 2,461$)	66.1 (10.6)	44:56	33:67	6.3 (3.7)	43.5 (10.8)
Dutch (18–49 yr, $n = 3,996$)	No ($n = 1,506$)	Reference[a] ($n = 1,506$)	35.0 (8.8)	39:61	59:41	2.1 (1.4)	50.7 (7.2)
	Yes ($n = 2,490$)	Chronicity ($n = 2,005$)	36.8 (8.5)	38:62	56:44	5.2 (3.5)	47.9 (9.1)
		Time course ($n = 2,013$)	36.8 (8.8)	38:62	56:44	5.2 (3.5)	47.9 (9.1)
		Health care use ($n = 2,196$)	36.8 (8.5)	37:63	56:44	5.2 (3.5)	47.9 (9.1)
		Patient impact ($n = 1,799$)	36.8 (8.5)	39:61	55:45	5.2 (3.5)	47.9 (9.1)

Notes: Patients were categorized separately on the four scales classifying consultation for any one of 115 morbidities.
Physical health was measured by PCS (Physical Component Score) of the Short Form-12 questionnaire.
Abbreviations: SD, standard deviation; M, males; and F, females.
[a] The reference group consists of consulters for other morbidities.

consenters and non-consenters in relation to their physical health scores ($P = 0.07$). The Dutch participants were representative of their overall population in relation to age, gender, and socioeconomic deprivation [14].

5. Discussion

5.1. Study findings

We have developed four new scales to classify morbidity severity that can be applied to routinely collected general practice consultation data. Morbidity severity classified on four separate scales was associated with older age, female gender, socioeconomic deprivation, and poor physical health. Our study provides evidence that clinical morbidity across the spectrum seen in general practice can be ordered by severity, using clinical constructs developed by GPs. This framework is a measure of physical health, and it offers the potential to identify health care needs in a standard way in primary care, using readily available data that accumulate personal health history over time.

The four scales and the order of severity categories within them had been determined on the basis of a consensus procedure involving GPs. The strength of the observed associations generally increased across the predetermined categories within all scales, and this provides empirical justification for the hierarchy of severity categories developed by the clinician consensus. The only exception was that the ''progressive'' category was more strongly linked with poor physical health than the ''permanent'' category. It is reasonable to consider that a progressive problem might have greater impact on health status than a stable but permanent condition, and both are features of higher impact morbidities [27,28]. The practical implication is that the two categories need to be reordered on the severity scale.

When the same validation tests used in the English setting were applied to a different general practice population—the national Dutch population aged 18 years and over—very similar trends were found for all sociodemographic characteristics and for the association between severity and poor physical health, in both the older and younger Dutch consulting populations. However, in younger Dutch consulters the associations were weaker, which probably reflects the overall better physical status of the young in comparison with the old consulters.

Current attempts at developing measures of morbidity severity have not used a peer-review led process [7,8,9], are not easily applied to routinely collected data [8], or have used a limited number of morbidities [9]. Most have not been validated in different general practice populations. We have now taken this forward in the following ways: we used clinical consensus to derive a severity classification that relates to daily practice, introduced four clear standardized severity scales that can be applied to routine consultation data, and validated this classification in general practice populations derived from two different countries. Previous studies have compared broad morbidity groups, for example, any cardiovascular disorder compared to any musculoskeletal disorder, but have not focused on the range of morbidities that can occur even within each broad morbidity group [29,30].

The potential usefulness of a morbidity severity classification as applied to routine consultations can be further extended to the study of multiple morbidities in the same individual. The scope of this analysis was validating tests of the classification within populations and between

Table 2
Distribution of the four separate morbidity severity scales by sociodemographic factors in English and Dutch consulters

Severity scales	Classifying categories	Age (yr)						Gender		Socioeconomic status	
		18–34 (%)	35–49 (%)	50–59 (%)	60–69 (%)	70–79 (%)	80+ (%)	Male (%)	Female (%)	Affluent (%)	Deprived (%)
Group	Reference[a]	–	–	21.2	14.1	13.7	10.7	17.2	14.8	17.8	13.4
		41.8	*34.9*	*23.7*	*17.5*	*14.0*	*12.6*	*28.6*	*28.3*	*32.5*	*24.6*
Chronicity	Acute	–	–	22.9	19.9	19.6	19.9	21.2	20.4	21.6	19.7
		30.1	*28.9*	*23.1*	*19.7*	*16.6*	*13.4*	*23.9*	*25.2*	*26.0*	*23.7*
	Acute-on-chronic	–	–	12.2	10.2	7.7	7.3	8.9	10.7	9.8	9.9
		14.0	*13.5*	*11.4*	*8.8*	*6.0*	*6.7*	*11.0*	*11.7*	*12.5*	*10.5*
	Chronic	–	–	27.1	43.1	49.0	50.5	39.3	41.4	37.5	44.1
		3.0	*9.4*	*27.9*	*43.3*	*53.8*	*54.5*	*24.8*	*22.2*	*16.6*	*29.4*
	Life-threatening	–	–	0	0.1	0.2	0	0.1	0.2	0.1	0.1
		0.3	*0.3*	*0.6*	*0.9*	*1.4*	*2.8*	*0.9*	*0.5*	*0.6*	*0.8*
Time course	One-off	–	–	12.1	10.4	8.2	8.3	9.6	10.6	10.2	10.0
		17.9	*16.1*	*11.5*	*9.5*	*6.4*	*3.9*	*12.0*	*13.8*	*14.2*	*12.0*
	Recurrent	–	–	28.0	23.1	20.4	19.0	23.8	23.2	23.5	23.4
		26.7	*28.4*	*24.7*	*23.4*	*19.1*	*18.7*	*24.6*	*25.5*	*25.5*	*24.9*
	Progressive	–	–	7.2	11.7	15.7	25.4	12.6	12.9	12.5	13.1
		1.2	*2.6*	*7.2*	*11.9*	*17.0*	*23.7*	*8.7*	*5.9*	*4.5*	*9.2*
	Permanent	–	–	17.3	28.0	30.9	26.6	23.3	26.9	23.2	27.9
		1.2	*5.7*	*18.9*	*28.4*	*36.2*	*35.5*	*14.6*	*15.6*	*11.7*	*18.5*
Health care use	Low	–	–	25.6	22.6	21.9	21.7	22.6	23.8	24.0	22.3
		38.3	*35.7*	*28.1*	*26.9*	*23.4*	*20.7*	*30.7*	*32.1*	*32.8*	*30.3*
	Medium	–	–	35.9	44.7	48.0	51.1	40.8	45.6	41.6	45.9
		13.3	*21.1*	*33.7*	*41.6*	*47.5*	*49.4*	*27.2*	*30.2*	*25.1*	*32.9*
	High	–	–	0.2	0.3	0.3	1.2	0.5	0.3	0.3	0.5
		0.2	*0.2*	*1.6*	*3.0*	*5.2*	*5.0*	*2.6*	*1.0*	*1.2*	*2.0*
Patient impact	Low	–	–	11.3	11.2	9.9	9.9	11.0	10.7	11.4	10.0
		11.7	*9.6*	*8.4*	*10.0*	*6.5*	*6.4*	*10.5*	*8.6*	*9.6*	*9.2*
	Medium	–	–	39.6	34.2	30.3	27.0	32.5	35.4	34.6	33.5
		28.4	*33.8*	*39.6*	*32.3*	*28.5*	*21.5*	*29.4*	*34.5*	*32.7*	*31.9*
	High	–	–	10.9	22.1	28.2	37.3	22.7	21.2	19.5	24.7
		1.2	*4.1*	*13.0*	*23.2*	*34.7*	*45.8*	*16.1*	*11.1*	*7.5*	*18.6*
Total numbers	English consulters	–	–	2,821	2,857	2,362	963	3,999	5,004	4,993	3,986
	Dutch consulters	*1,626*	*2,370*	*1,446*	*1,082*	*871*	*358*	*3,185*	*4,568*	*3,451*	*3,903*

Notes: Nonitalic figures relate to English consulters (50 years and over) and italic figures to Dutch consulters (18 years and over), and consulters were categorized separately on the four scales.

Percentage figures:

(i) are proportions of patients with the respective demographic characteristic who consulted at least once for a specified morbidity severity category in the 12-month period, for which the denominator populations are given in the bottom row,

(ii) are consulters for any one of 115 classified morbidities, and the [a]reference group is all other morbidity consulters, and

(iii) add up to less than 100% for each scale because some patients consulted for one of 115 morbidities, but morbidity had not been classified.

populations. Further work will use the classification to investigate the complexity of comorbidity in index conditions and the role of multimorbidity in influencing health and health care use. A further challenging potential is the use of the classification in actual decision-making process made by clinicians in consultations.

5.2. Design issues

The four scales were developed by clinicians as different ways of classifying morbidity, but there are obvious links between them, such as that between ''acute'' (chronicity) and ''one-off'' (time course) categories. However, only around a quarter of the morbidity set had been classified on all four scales, which means each scale classified different individuals. This suggests that the scales are discriminating between different aspects of morbidity severity, which justifies their separation in the analysis.

We chose physical health status as one measure of the ''generic'' health status, because the severity scales were developed with an overall ''health'' concept in mind, but other components of the health status, such as psychological status and specific subcomponents of the health surveys may offer alternative validation tests. Two of the scales (chronicity and time course) are about the ''nature of morbidity'' and imply a link to health status, that is, ''chronic'' is generally taken to mean worse health status than ''acute.'' These descriptive categories are used widely in clinical practice and research, and yet there has been little attempt to examine these concepts formally [27,31]. The health care use scale gives an overall picture in relation to different morbidities, but actual costs may be determined

Table 3
Associations between the separate morbidity severity scales and physical health

		English consulters (50+ yr)			Dutch consulters (50+ yr)			Dutch consulters (18–49 yr)		
Severity scales	Classifying categories	Health Good: Poor	Crude OR (95% CI)	OR[a] (95% CI)	Health Good: Poor	Crude OR (95% CI)	OR[a] (95% CI)	Health Good: Poor	Crude OR (95% CI)	OR[a] (95% CI)
Group[b]	Reference	840:552	1.0	1.0	522:177	1.0	1.0	1,057:449	1.0	1.0
Chronicity	Acute	918:901	1.5 (1.3–1.7)	1.4 (1.2–1.6)	453:287	1.9 (1.5–2.3)	1.9 (1.5–2.4)	726:448	1.5 (1.2–1.7)	1.5 (1.2–1.7)
	Acute-on-chronic	394:474	1.8 (1.5–2.2)	1.8 (1.5–2.2)	182:154	2.5 (1.9–3.3)	2.6 (1.9–3.4)	299:250	2.0 (1.6–2.4)	2.0 (1.6–2.4)
	Chronic	1,300:2,257	2.6 (2.3–3.0)	2.2 (1.9–2.5)	776:760	2.9 (2.4–3.5)	2.5 (2.0–3.1)	127:143	2.7 (2.0–3.4)	2.5 (1.9–3.3)
	Life-threatening	2:8	6.1 (1.3–28.8)	5.4 (1.1–25.6)	20:21	3.1 (1.6–5.8)	2.2 (1.1–4.4)	7:5	1.7 (0.5–5.3)	1.7 (0.5–5.3)
Time course	One-off	464:424	1.4 (1.2–1.6)	1.3 (1.1–1.6)	209:131	1.8 (1.4–2.4)	1.8 (1.4–2.5)	409:263	1.5 (1.3–1.8)	1.6 (1.3–1.9)
	Recurrent	972:1,085	1.7 (1.5–2.0)	1.6 (1.4–1.9)	494:349	2.1 (1.7–2.6)	2.0 (1.6–2.6)	633:474	1.8 (1.5–2.1)	1.7 (1.4–2.0)
	Progressive	284:835	4.5 (3.8–5.3)	3.5 (2.9–4.2)	185:281	4.5 (3.5–5.8)	3.9 (3.0–5.1)	35:45	3.0 (1.9–4.8)	2.7 (1.7–4.3)
	Permanent	914:1,311	2.2 (1.9–2.5)	1.8 (1.6–2.1)	551:471	2.5 (2.0–3.1)	2.1 (1.7–2.7)	78:76	2.3 (1.6–3.2)	2.2 (1.5–3.1)
Health care use	Low	1,019:1,023	1.5 (1.3–1.8)	1.4 (1.2–1.6)	583:393	2.0 (1.6–2.5)	1.9 (1.5–2.4)	916:553	1.4 (1.2–1.7)	1.4 (1.2–1.7)
	Medium	1,466:2,358	2.4 (2.2–2.8)	2.1 (1.8–2.4)	792:737	2.7 (2.3–3.3)	2.4 (2.0–3.0)	360:358	2.3 (1.9–2.8)	2.2 (1.9–2.7)
	High	9:23	3.9 (1.8–8.5)	2.8 (1.2–6.4)	55:64	3.4 (2.3–5.1)	3.4 (2.2–5.3)	6:3	1.2 (0.3–4.7)	1.1 (0.3–4.4)
Patient impact	Low	529:417	1.2 (1.0–1.4)	1.1 (0.9–1.3)	216:93	1.3 (0.9–1.7)	1.2 (0.9–1.7)	271:147	1.3 (1.0–1.6)	1.3 (1.0–1.6)
	Medium	1,357:1,654	1.9 (1.6–2.1)	1.8 (1.6–2.0)	708:539	2.2 (1.8–2.8)	2.1 (1.7–2.6)	723:540	1.8 (1.5–2.1)	1.7 (1.5–2.0)
	High	459:1,451	4.8 (4.1–5.6)	3.7 (3.2–4.3)	278:527	4.1 (3.3–5.1)	3.5 (2.8–4.4)	49:69	3.3 (2.3–4.9)	3.2 (2.1–4.7)

Notes: Patients were categorized separately on the four scales for consultation for any one of 115 morbidities.
Physical health was measured by PCS (Physical Component Score) of the Short Form-12 questionnaire.
[a] Odds ratio (OR) adjusted for age, gender, and deprivation.
[b] Reference group consists of consulters for other morbidities.

by the effectiveness of available treatments. The patient impact scale relates clinician judgments about the likely effect on activities of daily living, but patients' perceptions of their morbidity severity may differ from their clinician [32]. Our morbidity list is still only a sample of all those seen in general practice, and further work might lead to all morbidities being classifiable on our severity scales. Our approach to severity measurement is also distinct to other methods that use actual patient assessment to define severity of morbidity. The main advantage of our approach is that it can be readily applied to routinely collected individual morbidity data, providing an indicator of physical status and is useful to assess and compare on a group level, without requiring formal time-consuming assessment of the patient. However, our classification still requires additional work, for example, validation against other outcomes and comparison with other scales of severity.

The associations between morbidity severity and poor physical health were also adjusted for age, gender, and deprivation to take into account the differences between the English and Dutch consulting populations. We purposely chose the undefined morbidity group as reference, as opposed to nonconsulters, as the intention was to compare groups drawn from the same population, that is, people who consult for a morbidity. Participants had characteristics associated with increased likelihood of general practice attendance, for example, older age groups and females, but these were also the same groups who were less likely to consent. Any under- or overestimates of consultation prevalence arising from this is unlikely to affect the internal comparison of case and reference groups.

Individual diagnoses in routinely collected morbidity data may be subject to misclassification, especially because general practice is the first point of contact for a range of nonspecific symptoms. However, our approach allocated patients into exclusive and most severe category for many different morbidities, thus reducing the role of misclassification. Furthermore, multiple consultations for different morbidities over a 12-month period are likely to provide an accurate reflection of an individual's general health status over time. Previous studies have also validated this use of general practice data, in particular for chronic diseases [33,34].

5.3. Conclusions

We have developed a practical approach for classifying general practice populations by morbidity severity using routine consultations, which is suitable for use in clinical practice. It has potential to be incorporated into actual consultation data, with appropriate software, to provide readily available and dynamic measure of health status and consequently health needs in real-time clinical practice. Current research has already noted, for example, that older patients with complex chronic diseases may require a case-management approach that focuses on overall health care needs [35,36]. Our framework offers the potential to identify and target a similar but broader group of patients in primary care consulting populations. We are using it now to

examine the role of multimorbidity in general practice, and further research will address its value in case management.

Acknowledgments

U.T.K. has been supported by a Medical Research Council (UK) Training Fellowship in Health Services Research (G106/1035). Other funding support was by Claire Wand Fund, North Staffordshire Primary Care Research Consortium, MRC Programme grant (G9900220), and NHS R&D funding. We wish to thank all patients, North Staffordshire GP Consortium group, GPs, and the Primary Care Network and Surveys teams at Keele University, and staff at EMGO (Amsterdam) and NIVEL (Utrecht).

References

[1] Bowling A. Research methods in health. Buckingham: Open University Press; 1997.
[2] McCormick A, Fleming D, Charlton J. Morbidity Statistics from General Practice: Fourth National Study 1991−92. Series MB5 No. 3. Office of Population Censuses and Surveys. London: HMSO; 1995.
[3] Kadam UT, Croft P, Lewis M. Use of a cross-sectional survey to estimate outcome of health care: the example of anxiety and depression. J Clin Epidemiol 2001;54:1112−9.
[4] Hay EM, Mullis R, Lewis M, Vohora K, Main CJ, Watson P, et al. Comparison of physical treatments versus a brief pain-management programme for back pain in primary care: a randomised clinical trial in physiotherapy practice. Lancet 2005;365:2024−30.
[5] UK Beam Trial Team. United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: effectiveness of physical treatments for back pain in primary care. BMJ 2004;329:1377.
[6] Neal RD, Heywood PL, Morley S, Clayden AD, Dowell AC. Frequency of patients' consulting in general practice and workload generated by frequent attenders: comparisons between practices. Br J Gen Pract 1998;48:895−8.
[7] Weiner JP, Starfield BH, Steinwachs DM, Mumford LM. Development and application of a population-oriented measure of ambulatory care case-mix. Med Care 1991;29:452−72.
[8] Parkerson GR Jr, Michener JL, Yarnall KS, Hammond WE. Duke Case-Mix System for ambulatory health care. J Clin Epidemiol 1997;50:1385−94.
[9] van den Bosch WJ, Huygen FJ, van den Hoogen HJ, van Weel C. Morbidity in early childhood: differences between girls and boys under 10 years old. Br J Gen Pract 1992;42:366−9.
[10] Royal College of General Practitioners, Office of Population Censuses and Surveys, and Department of Health. Morbidity statistics from general practice 1981−82: third national study (MSGP3). Series MB5 No.1; London; HMSO: 1986.
[11] Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use. 2nd ed. Oxford: OUP; 2001.
[12] Deckers JGM, Schellevis FG. Health information from primary care. Final report, December 1, 2001−March 31, 2004. Utrecht: NIVEL; 2004.
[13] Harding A, Stuart-Buttle C. The development and role of the Read codes. J Ahima 1998;69:34−8.
[14] Kadam UT, Jordan K, Croft PR. A comparison of two consensus methods in a single professional group showed the same outcomes. J Clin Epidemiol 2006;59:1169−73.
[15] Westert GP, Schellevis FG, Bakker DH de, Groenewegen PP, Bensing JM, Zee J van der. Monitoring health inequalities through General Practice: the Second Dutch National Survey of General Practice. Eur J Public Health 2005;15:59−65.
[16] Porcheret M, Hughes R, Evans D, Jordan K, Whitehurst T, Ogden H, et al. Data quality of general practice electronic health records: the impact of a program of assessments, feedback, and training. J Am Med Inform Assoc 2004;11:78−86.
[17] Bentsen BG. International classification of primary care. Scand J Prim Health Care 1986;4:43−50.
[18] Jinks C, Jordan K, Ong BN, Croft P. A brief screening tool for knee pain in primary care (KNEST). 2. Results from a survey in the general population aged 50 and over. Rheumatology 2004;43:55−61.
[19] Thomas E, Wilkie R, Peat G, Hill S, Dziedzic K, Croft P. The North Staffordshire Osteoarthritis Project—NorStOP: prospective, 3-year study of the epidemiology and management of clinical osteoarthritis in a general population of older adults. BMC Musculoskelet Disord 2004;5:2.
[20] Townsend P, Phillimore P, Beattie A. Health and deprivation: Inequality and the north. London: Croom Helm; 1988.
[21] Ware JE Jr, Sherbourne CD. The MOS 36-item short form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473−83.
[22] Jenkinson C, Layte R, Jenkinson D, Lawrence K, Petersen S, Paice C, et al. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? J Public Health Med 1997;19:179−86.
[23] Kosinski M, Bayliss MM, Bjorner JB, Ware JE Jr. Improving estimates of SF-36 health survey scores for respondents with missing data. Medical Outcomes Trust Monitor 2000;5:8−10.
[24] Brayne C, Matthews FE, McGee MA, Jagger C. Health and ill-health in the older population in England and Wales. The Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). Age Ageing 2001;30:53−62.
[25] Vedsted P, Christensen MB. Frequent attenders in general practice care: a literature review with special reference to methodological considerations. Public Health 2005;119:118−37.
[26] McNiece R, Majeed A. Socioeconomic differences in general practice consultation rates in patients aged 65 and over: prospective cohort study. BMJ 1999;319:26−8.
[27] O'Halloran J, Miller GC, Britt H. Defining chronic conditions for primary care with ICPC-2. Fam Pract 2004;21:381−6.
[28] Wagner EH, Groves T. Care for chronic diseases. BMJ 2002;325:913−4.
[29] Sprangers MA, de Regt EB, Andries F, van Agt HM, Bijl RV, de Boer JB, et al. Which chronic conditions are associated with better or poorer quality of life? J Clin Epidemiol 2000;53:895−907.
[30] Wilder CS. Report No. 15. Acute Conditions: Incidence and associated disability: United States, July 1963−June 1964. National Centre for Health Statistics. Available at. http://www.cdc.gov/nchs/products/pubs/pubd/series. Accessed December 23, 2005.
[31] Kuller LH. Relationship between acute and chronic disease epidemiology. Yale J Biol Med 1987;60:363−77.
[32] Memel DS, Kirwan JR, Sharp DJ, Hehir M. General practitioners miss disability and anxiety as well as depression in their patients with osteoarthritis. Br J Gen Pract 2000;50:645−8.
[33] Jordan K, Porcheret M, Croft P. Quality of morbidity coding in general practice computerized medical records: a systematic review. Fam Pract 2004;21:396−412.
[34] Jick H, Jick SS, Derby LE. Validation of information recorded on general practitioner based computerised data resource in the United Kingdom. BMJ 1991;302:766−8.
[35] Von Korff M, Gruman J, Schaefer J, Curry SJ, Wagner EH. Collaborative management of chronic illness. Ann Intern Med 1997;127:1097−102.
[36] Hutt R, Rosen R, McCauley J. Case managing long-term conditions: What impact does it have in the treatment of older people?. London: King's Fund; 2004.

Examples of morbidities classified on severity scales

	Severity scales			
Clinical codes	Chronicity	Time course	Health care use	Patient impact
Infectious gastroenteritis	Acute	One-off	Low	Undefined*
Herpes zoster	Acute	One-off	Low	Medium
Dermatophytosis of foot	Acute	Undefined	Low	Medium
Candidiasis	Acute	One-off	Low	Undefined
Hypothyroidism	Chronic	Permanent	Low	Undefined
Diabetes mellitus	Chronic	Undefined	Undefined	High
Pure hypercholesterolaemia	Chronic	Permanent	Undefined	Low
Gouty arthropathy	Acute-on-Chronic	Recurrent	Undefined	Medium
Obesity	Chronic	Undefined	Undefined	Medium
Anxiety states	Acute-on-Chronic	Recurrent	Medium	Undefined
Nonorganic sleep disorders	Undefined	Recurrent	Low	Medium
Depressive disorder	Undefined	Recurrent	Medium	Undefined
Senile cataract	Chronic	Progressive	Medium	High
Conjunctivitis	Acute	One-off	Low	Low
Wax in ear	Acute	Recurrent	Low	Low
Labyrinthitis	Acute	One-off	Low	Medium
High blood pressure	Chronic	Permanent	Medium	Undefined
Angina pectoris	Undefined	Undefined	Undefined	High
Atrial fibrillation	Chronic	Permanent	Medium	Medium
Congestive heart failure	Undefined	Progressive	Undefined	High
Haemorrhoids	Acute-on-Chronic	Recurrent	Low	Undefined
Nasopharyngitis	Acute	One-off	Low	Low
Sinusitis	Acute	One-off	Low	Undefined
Sore throat	acute	One-off	Low	Low
Tracheitis	Acute	One-off	Low	Low
Upper respiratory tract infection	Acute	One-off	Low	Low
Bronchitis	Acute	Undefined	Undefined	Undefined
Allergic rhinitis	Acute-on-Chronic	Recurrent	Low	Undefined
Influenza	Acute	One-off	Low	Undefined
Chronic bronchitis	Undefined	Progressive	Medium	Undefined
Asthma	Acute-on-Chronic	Recurrent	Medium	Medium
Esophagitis	Acute-on-Chronic	Recurrent	Medium	Medium
Indigestion	Undefined	Recurrent	Undefined	Medium
Constipation—functional	Undefined	Recurrent	Low	Medium
Urinary tract infection	Acute	One-off	Low	Undefined
Menopausal or female climacteric state	Undefined	Undefined	Undefined	Medium
Atopic dermatitis	Acute-on-Chronic	Recurrent	Undefined	Medium
Eczema	Undefined	Recurrent	Low	Medium
Pruritus	Acute	Undefined	Low	Undefined
Seborrheic wart	Undefined	Undefined	Low	Low
Generalized osteoarthritis	Chronic	Progressive	Medium	High
Knee joint pain	Undefined	Undefined	Undefined	Medium
Cervical spondylosis	Undefined	Undefined	Undefined	Medium
Back pain	Undefined	Recurrent	Undefined	Medium
Rotator cuff shoulder syndrome	Undefined	Undefined	Medium	Medium
Tenosynovitis	Acute	Undefined	Undefined	Medium
Drug adverse effects	Acute	One-off	Low	Undefined

Note: Undefined* means that morbidity was not classified by severity by GP consensus; examples are based on the most prevalent morbidities from a national English sample (MSGP4) (McCormick et al., 1995 [2]).